

P1 188348

REC'D 14 JAN 2000

WIPO

PCT

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

January 10, 2000

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM
THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK
OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT
APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A
FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: 60/108,291
FILING DATE: November 13, 1998
PCT APPLICATION NUMBER: PCT/US99/26562

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS



H. Phillips
H. PHILLIPS
Certifying Officer

11/13/98



618 U.S. PTO

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR §1.53(b)(2).

Docket No. 98,675		Type a plus sign (+) inside this box: +	
INVENTOR(S)/APPLICANTS(S)			
LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (City and either state or foreign country)
Boyce	Keith		Wexford, PA
McKenna	Brian		Pittsburgh, PA
TITLE OF THE INVENTION (280 character maximum)			
System and method to achieve integrated presentation of summary statistics, detail measurements, and raw images while maintaining efficient use of storage media through separation of the data in an integrated data management and presentation system			
CORRESPONDENCE ADDRESS			
McDonnell Boehnen Hulbert & Berghoff 300 South Wacker Drive, Chicago			
STATE	Illinois	ZIP CODE	60606 COUNTRY U.S.A.
ENCLOSED APPLICATION PARTS (check all that apply)			
<input checked="" type="checkbox"/> Specification Number of Pages <u>12</u> <input type="checkbox"/> Small Entity Statement <input checked="" type="checkbox"/> Drawing(s) Number of Sheets <u>within specification</u> <input type="checkbox"/> Other (specify):			
METHOD OF PAYMENT FOR THIS PROVISIONAL APPLICATION FOR PATENT			
XX A check or money order is enclosed to cover the Provisional Filing Fee. The Commissioner is hereby authorized to charge filing fees and credit Deposit Account Number: 13-2490.		PROVISIONAL APPLICATION FOR PATENT FILING FEE AMOUNT (\$)	75.00

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.
☒ No. ☐ Yes, the name of the U.S. Government agency and the Government contract number are: _____

Respectfully submitted,

SIGNATURE: David HarperDate: 11/13/98TYPED or PRINTED NAME David HarperREG. NO. 42,636

Additional inventors are being named on separately numbered sheets attached hereto.

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

Burden Hour Statement: This form is estimated to take 2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Office of Assistance Quality and Enhancement Division, Patent and Trademark Office, Washington, D.C. 20231, and to the Office of Information and Regulatory Affairs, Office of Management and Budget, Executive Order 12812, Washington, D.C. 20503.

11/13/98



518 U.S. PTO

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
(Case No. 98,675)

In re Application of:)	
)	
Boyce and McKenna)	
)	Art Unit: To be Assigned
Serial No.: To be Assigned)	
)	Examiner: To be Assigned
Filed: To be Assigned)	
)	
For: System and method to achieve integrated presentation)	
of summary statistics, detail measurements, and raw)	
images while maintaining efficient use of storage media)	
through separation of the data in an integrated data)	
management and presentation system)	

Asst. Commissioner for Patents
BOX PROVISIONAL APPLICATION
Washington, D.C. 20231

TRANSMITTAL LETTER

Sir:

1. We are transmitting herewith the attached papers for the above identified new provisional patent application:

- ☒ Patent Specification (11 pages, including cover sheet, claims, and abstract)
- ☒ Drawings (within specification)
- ☒ Return Postcard
- ☒ Other: Provisional Application Cover Sheet

2. ☒ A check in the amount of \$75.00 is enclosed for the Filing Fee.

☐ Please charge the total filing fee of \$ to our Deposit Account No. 13-2490. A duplicate copy of this sheet is enclosed.

3. **GENERAL AUTHORIZATION TO CHARGE OR CREDIT FEES:** Please charge any additional fees or credit overpayment to Deposit Account No. 13-2490. A duplicate copy of this sheet is enclosed.
4. **CERTIFICATE OF MAILING BY "EXPRESS MAIL" UNDER 37 CFR § 1.10:** The undersigned hereby certifies that this Transmittal Letter and the paper, as described in paragraph 1 hereinabove, are being deposited with the United States Postal Service with sufficient postage as "Express Mail Post Office to Addressee" in an envelope addressed to: Asst. Commissioner for Patents, Box New Application, Washington, D.C. 20231, on this 13th day of November, 1998. Express Mail No. EM004663800US

By: Mark Rodenburg

Cellomics™ Store

Cellomics Inc.

1. Product Overview

1.1. Overview

Cellomics™ Store Version 1.0 is the first release of the Cellomics™ Store product. It is a client server system designed to store and provide cross-plate analysis of plate-screening data from one or many ArrayScan systems. It has three primary components, The Cellomics™ Store database, the Cellomics™ Store Client, and the Cellomics™ Store Archive. The primary requirements for Release 1.0 of Cellomics™ Store are to provide data management and archival functions for the detailed data associated with one or many ArrayScan systems, and to provide integrated data visualizations of the well data, cell data, and cell images on a client computer.

1.2. Background

1) Provide data management and archival functions for ArrayScan data

Cellomics™ Store will focus on data management and archival in the first release. With the separation of image files from data files the 2 types can be managed according to different policies and procedures. In the Cellomics™ Store 1.0 timeframe, the largest data management challenge is with image data, but detailed field and cell data can also be voluminous. Image files will first be migrated to server disk storage, then to an optical juke box, and eventually either deleted or copied to off-line storage if necessary. ArrayScan data will also be centrally managed and archived through an archiving system for plate database files. The goal will be to automate as much of this migration as possible based on available storage.

2) Allow on-line viewing of detailed Arrayscan data from remote PC's that are not full ArrayScan stations

Cellomics™ Store will provide the ability for remote workstations that do not have the ArrayScan hardware to review and analyze data from other ArrayScan stations after a scan is completed and migrated to Cellomics™ Store.

3) Create a data warehouse of plate screening data for cross-plate analysis.

Cellomics™ Store will provide an integrated client-server database of plate and well level data. This integrated database will be the basis for cross-plate analysis. References to the image files will be stored with the data so that an integrated presentation of data and images can be shown in the user interface. In addition to storing plate data, Cellomics™ Store will provide cross-plate analysis functions to validate ArrayScan plate screening runs before data is "published" to corporate drug discovery databases.

1.3. Overall Requirements Summary

Given the vast amount of image data that is generated by ArrayScan systems, there is a need to provide a central repository of that data where it can be analyzed, managed and archived. The current ArrayScan system can only review a single plate at a time. Cellomics™ Store will enable a workgroup to share in the analysis of ArrayScan screening data.

2. Functionality

2.1. Overview

Cellomics™ Store is a data management system for storing and managing the images and data generated from ArrayScan Cell-Based Screening systems. Summary data from the ArrayScan plate screens (plate and well level) is automatically entered into the Cellomics™ Store Database for cross-plate analysis and validation. The detailed cellular measurements database file and image files are automatically copied to a central Cellomics™ Store shared directory. The plate database files and image files will be archived according to user defined Hierarchical Storage Management (MSM) policies.

Cellomics™ Store is designed as a client server system that can connect scientists in screening workgroups to the data from one or more Arrayscan systems. Cellomics™ Store makes the data from all of the Arrayscan stations available to all of the scientists within the workgroup. Users can use the Cellomics™ Store Client application to access the detailed screening data and images from any ArrayScan within the workgroup, and to perform cross-plate analysis, including standard graphs, and reports of plate-screening data.

3. Configuration and Interfaces

3.1. Overview

As much as possible, Cellomics™ Store will use industry standard tools. The Cellomics™ Store product will try to "bundle" other products together when it makes programming and economical sense to do so (e.g. using a data management product for interfacing to the Jukebox, rather than writing a new one). The target environment will be Windows/NT 4.0 or Windows/95 and above for the Cellomics™ Store Client applications, and Windows/NT 4.0 and above for the Cellomics™ Store Server. The Cellomics™ Store client will not be supported in the Windows 3.x environment.

3.2. Computer Platform

3.2.1. Hardware Platform - Cellomics™ Store Server

For Cellomics™ Store 1.0 sites with up to 3 ArrayScan systems, the following configuration is recommended:

- Pentium 266 MHz. or above processors with 256 MB of RAM, PCI bus slot, CD-ROM Drive, and SCSI controller (e.g. DIGITAL Server 5200 or similar).
- Three or more separate disk volumes of at least 9 GB each
- Video card and color monitor that support 1024 X 768 resolution
- Optical Jukebox mass storage subsystem with at least 300 GB capacity (HP SureStore P 600 FX or similar)
- Industry standard 10/100 Ethernet network card

3.2.2. Primary Operating System Environment - Cellomics™ Store Server

- Microsoft Windows/NT Server 4.0
- Microsoft SQL Server Version 6.5 or higher
- Seagate Software Storage Migrator for Windows/NT

3.2.3. Primary Environment - Cellomics™ Store Client

- Windows 95/98 or Windows NT 4.0 or higher
- Monitor that supports video resolution of 1024 x 768 or higher
- Hardware capable of running Windows/NT or Windows/95
- Industry standard 10/100 Ethernet network card

3.3. Other Hardware Requirements

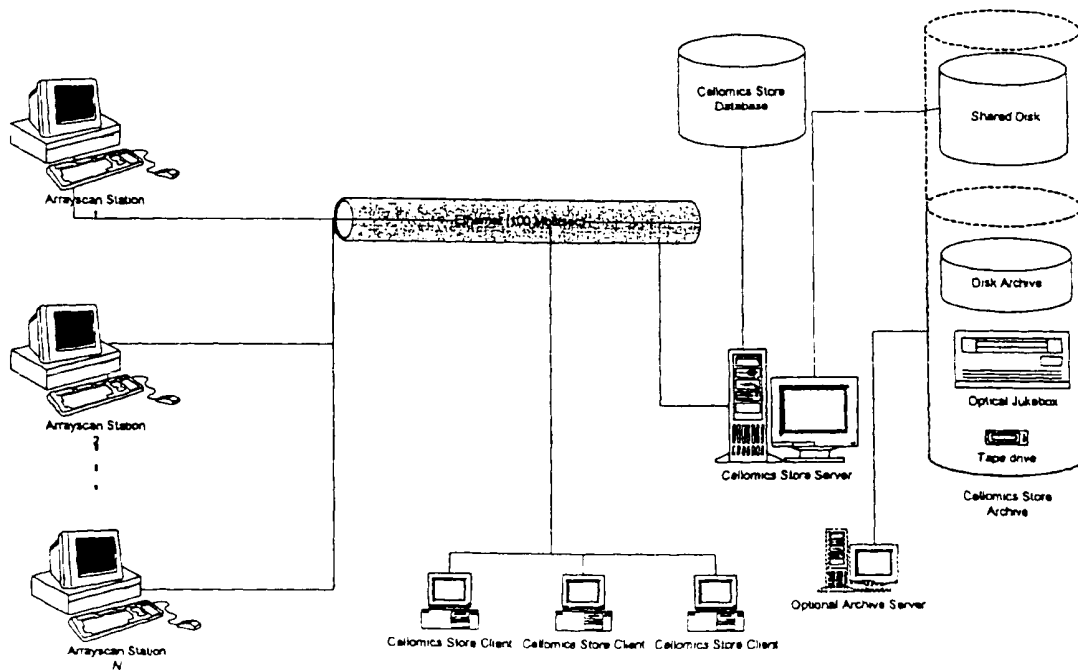
As an additional level of hierarchical storage for the Arrayscan detailed data, customers may want to add an automated Tape Library. An alternative would be to use existing Tape Libraries within the MIS organization since Seagate's Storage Migrator can be integrated with IBM's ADSM product line. For those users who want to maintain their own tape library within their workgroup we would recommend a DLT Tape library subsystem with at least 300 GB capacity.

4. Architecture

4.1. Overview

The primary components of the Cellomics™ Store System are

- The Cellomics™ Store Server
- The Cellomics™ Store Database
- The Cellomics™ Store Archive
- The Cellomics™ Store Client



5.2. Cellomics™ Store Server

The Cellomics™ Store Server is the platform that controls all of the Cellomics™ Store server based functions plus the underlying Database Management System (DBMS).

5.3. Cellomics™ Store Database

The Cellomics™ Store Database is a multi-user database that stores the summary data from multiple ArrayScan Systems. It uses standard relational database tools and structures for all of its functions. The Data Model is defined in the Cellomics™ Store Detailed Design document.

5.4. Cellomics™ Store Archive

The Cellomics™ Store Archive is a library of ArrayScan image and database files stored on a high volume (300-600 GB) optical jukebox. Cellomics™ Store will use Hierarchical Storage Management (HSM) techniques to automatically manage the disk space of ArrayScan stations.

Cellomics™ Store will support two modes of image file and plate database file archiving. The first mode will be to retain the files on the ArrayScan station where they were originally written. Cellomics™ Store's archiving functions will automatically manage the free space on the ArrayScan disks to move files to the archival storage systems. To the end user the files will appear to be in the same directory where they were originally stored. The files may actually be stored on the server's disk, on an optical jukebox, or in a DLT library. The other mode for Cellomics™ Store archiving will be to move/copy the plate database and images to a central NT file server. The Cellomics™ Store archival functions will then manage this shared server using the same HSM techniques.

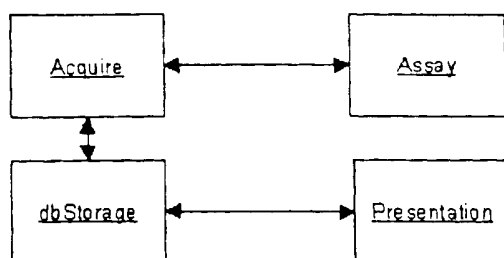
5.5. Cellomics™ Store Client

The Cellomics™ Store Client is a desktop resident program that provides the user interface to the Cellomics™ Store Database and Cellomics™ Store managed plate and image data. The application will focus on cross-plate display functions. It uses a GUI to lead the scientist through standard analyses, and also supports custom and ad-hoc query and viewing capabilities. It also supports data export into standard desktop tools such as spreadsheets, graphics packages, and word processors. The Cellomics™ Store client connects to the Cellomics™ Store Server through an ODBC connection.

5.6. Arrayscan Architecture

The ArrayScan software is divided into four functional groups or modules. In addition to dbStorage, the other modules are Acquisition, Assay and Presentation. The Acquisition module controls the robotic microscope and camera, acquires images and sends the images to the Assay Module. The Assay Module "reads" the images, creates graphic overlays, interprets the image and returns the new images and data extracted from the images back to the Acquisition Module. The Acquisition Module then passes the image and interpreted data to the dbStorage module. The dbStorage module saves the information in a combination of image files and relational databases. The Cellomics™ Store Client uses the dbStorage module to access the data and images for presentation and data analysis of the information acquired.

ArrayScan Modules



The ArrayScan system will be sold as either a stand-alone unit or as part of a Cellomics™ Store Bio-Informatics package that will handle, data storage and analysis in a multi-user, client-server environment. To support this functionality the dbStorage module must be configurable and be able to handle both modes of operation.

5.7. Volume of Data Assumptions

It is the primary task of the dbStorage module to format and store the large volume of data that the ArrayScan will produce. It must store data in a way that can be accessed via the presentation software, and allow for data mining and archiving. The raw data coming in from the ArrayScan can potentially be extremely large, on the order of 0.5 Gigabytes per hour per ArrayScan. Since an organization may have several ArrayScans, and the ArrayScans can operate 24 hours a day the possible data flow through the network may approach the physical limits for a conventional PC network.

To handle the various configurations of ArrayScan systems, the dbStorage module will be capable of storing the data on one machine or working in a client/server environment. A spooling function will provide a method of transferring the data into Cellomics™ Store to provide limitless storage ability. The dbStorage module will need to store data locally, spool data to a network server and archive data to CD-ROM or other data storage device. It must also be able to find and retrieve that information for single plates in Arrayscan mode.

The proposed design of dbStorage is based on several assumptions as to the volume of data that will be handled. If the volume of data is less than that assumed the system should be able to handle it without problem. If the volume is more, one of the following must happen:

- The volume must be decreased to the limits set
- Faster, more expensive hardware must be used
- The underlying database must be changed to a faster, high end product
- The design of the system must be changed
- Changes to the data to decrease volume (non loss-less compression, deletion, etc.)

The volume of data is determined by the amount of images and cell feature information accumulated for each plate scanned. The data can be stored on a local ArrayScan, on a Network Server Drive, or archived to CD-ROM. It is assumed that for the single user version, several plates worth of data will be stored locally and then manually archived to CD-ROM, or deleted. For the Cellomics™ Store Client/Server product, one or more ArrayScans will spool data to a network drive where they will be archived and/or deleted from the server by the Cellomics™ Store product

The ArrayScan performs Assay tests on Microtiter Plates. Each plate consists of 96 wells (384 well plates will be used in the next year or so). Each well is divided into fields, which represents the field of vision (zoom) of the microscope's camera. The number of fields per well will vary by camera with a physical maximum of 120 and a realistic maximum of 16. Each field will have between 1 and 6 images taken of it, each using a different light filter to capture a different wavelength of light. Each field will also find a number of cells to analyze. The number of cells will vary with 100 the upper limit and 10 the norm. For each cell the assay test will collect up to 10 features for each cell. From a data volume perspective, the data to be saved can be estimated by number of cell feature records stored and the number of images stored. The number of images can be calculated by the equation:

- (NUMBER OF WELLS X NUMBER OF FIELDS X IMAGES PER FIELD)

The current size of an image file is approximately 512 Kbytes of uncompressed data. Note that using standard image compression could lower the image size by about 50 percent and non-standard image compression could lower the size of the image file even more.

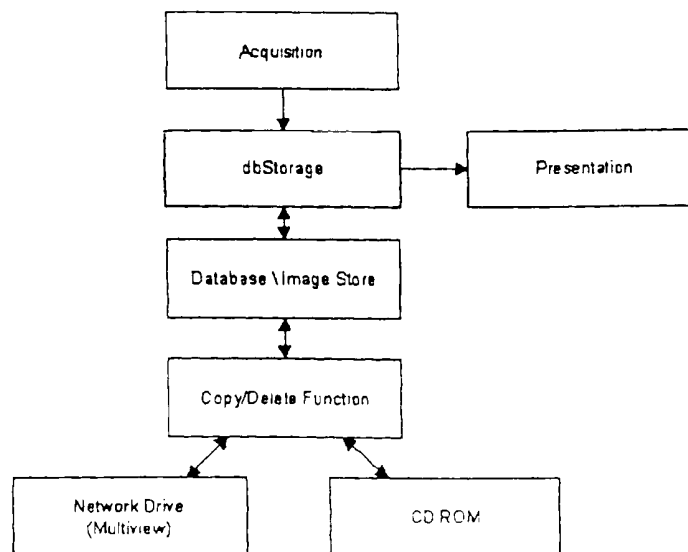
The number of cell feature records can be calculated by the equation:

- (NUMBER OF WELLS X NUMBER OF FIELDS X CELLS PER FIELD X FEATURES PER CELL)

The objective for the acquisition module is to be able to scan the "Realistic Goal" (below) type plate every 30 minutes and to be able to scan the "Worst Case Goal" (below) type plate over a longer time period:

Variable	Current Use	Realistic Goal	Worst Case Goal
WELLS PER PLATE	96	96	384
FIELDS PER WELL	3	4	16
IMAGES PER FIELD	2	3	4
CELLS PER FIELD	20	100	100
FEATURES PER CELL	10	10	10
SIZE OF EACH IMAGE	256K	512K	512K
Total number of images	576	1152	24576
Total size of images	144 Mbytes	576 Mbytes	12,288 Mbytes
Total rows in largest table	57600	384000	6144000
Size of database	5 Mbytes	34 Mbytes	540 Mbytes
Volume of data per hour	300 Mbytes	600 Mbytes	600 Mbytes
Hours use per day/days per week	8 / 5	20 / 5	24 / 7
Volume of data per day	2,400 Mbytes	12,000 Mbytes	14,400 Mbytes
Volume of data per week	12,000 Mbytes	60,000 Mbytes	100,800 Mbytes

Data Flow through ArrayScan

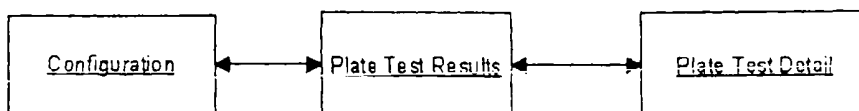


5.8. Image File Formats

Images acquired by the system will be stored as image files using a standard image format, such as TIF file format. The naming convention for the files will indicate the test it was from, but will rely on the database Image table to identify the image file with a particular section of the plate scan. Due the large volume of images, it is anticipated that a file compression option for image storage will be used. It is possible that future revisions of dbStorage will use image compression that induces image loss may be used to achieve larger compression ratios and increase the ability to archive data.

5.9. Database Format

ArrayScan Data by Functional Area



Data acquired by the ArrayScan will be stored in a relational database for use by the acquisition and presentation modules. There are three types of data to be maintained:

- Configuration data, which holds control parameters for the Acquire and Assay modules. This data will be stored in several tables all with fewer than 1,000 records and will not grow with additional scans.
- Plate scan results information, which holds information about each plate that has been run, or is scheduled to be run. This data can be stored in a few tables and while it will continue to grow with additional scans, the table with the most rows (WellFeatures) an upper limit of 1,000 records per scan can be handled without difficulty.
- Plate scan data represents the largest amount of data. This data will be stored in several related tables with the largest table (CellFeatures) growing up to 6,000,000 records for each plate run.

Because of the large volume of data, and because of the need for data mining and presentation, a highly normalized relational database structure is needed. The likely hood of 3rd party data presentation tools and the need for data management means that the use of a standard database system is desired. Three areas where this data will differ from many other relational database management systems is that:

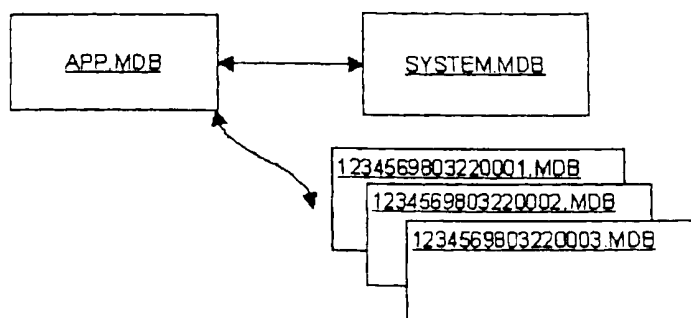
- The number of records is very large
- Almost all the data is write once/read only data
- A large percentage of this data (Plate Test Detail) will only be used for exception analysis.

Cost and scalability determine the choice of database system. Microsoft Access has been chosen for the ArrayScan system. For the Cellomics™ Store version it is anticipated that the back-end database will change to conform to the customer's IT strategy. For this reason the strategy of using an Access pass-through database to either other Access databases or to ODBC databases will be used. The reasons for choosing Access as the ArrayScan database:

- Can handle the anticipated size and records of "Worse Case Goal"
- Can be used as pass-through database to other ODBC databases
- Royalty-free distribution
- SQL compliant
- Industry acceptance
- Supports COM interface

It is important to note that the ArrayScan application will never have to handle tables larger than 6,000,000 records, and rarely look at views of over a few thousand records at one time. This is because there is no need for evaluation of plate detail data information across plates. Summary plate information can be stored in the plate, well, plate feature and well feature tables to be compared across plates. Detail information about individual cells will only need to be accessed with the context of evaluating within that one plate test. This allows the application can make use of the Microsoft Access feature of pass-through tables. The actual database the application is looking at that does not contain data, but is used as a pass-through to other databases that do contain the data. The database that contains the actual Plate Test Detail data is changed dynamically as the user moves from Plate Test to Plate Test. The database that contains the Configuration Data and Plate Test Results data is also linked, but is not changed - the application will always refer to the same data set for this information. The reason that this data is stored in a linked database is so that the actual database engine could be changed to another database type, such as Microsoft Sequel Server or Oracle, or the location of the database can change without the application having to be modified. The names for the databases are APP.MDB which will be the pass through MS Access database; SYSTEM.MDB which will contain the Configuration Data and Plate Test Results data. Since the Plate Test Detail data will be many databases, the database it self will be created when the plate record is created and will have a name that will be created by taking the Plate ID field value and adding ".MDB" to the end. (For example, a record in the plate table with an ID of "1234569803220001" will have it's data stored in a database with the name "1234569803220001.MDB".

ArrayScan Databases



5.10. Data Locations

The system will be configured either as a single computer system, or as part of the Cellomics™ Store Client/Server network. Both configurations must support archiving of data and deleting of local data. To support Cellomics™ Store, the ArrayScan must support the spooling of data off of the ArrayScan machine in a way to avoid network traffic problems effecting the acquisition of data from the ArrayScan.

For single computer systems the APP.MDB and the SYSTEM.MDB files will reside in a subdirectory of the executable program named "DATA". The default directory for this will be "C:\CELLOMICS\DATA". The created plate specific database files will reside in a subdirectory of the data directory named "PLATES". The default directory for this will be "C:\CELLOMICS\DATA\PLATES". The created image files will reside in a subdirectory of the data directory named "IMAGES". The default directory for this will be "C:\CELLOMICS\DATA\IMAGES".

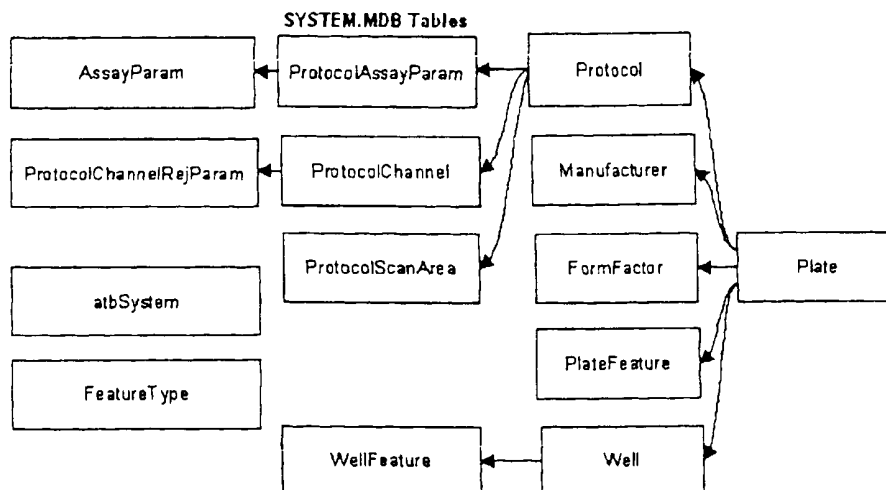
For the Cellomics™ Store version, the APP.MDB will reside on the Client computer in the application subdirectory "DATA". The default directory for this will be "C:\CELLOMICS\DATA". The SYSTEM.MDB will

reside on the Server computer in the application subdirectory "DATA". The default directory for this file will be "F:\CELLOMICS\DATA".

The plate specific database files will *FIRST* reside in a subdirectory of the data directory named "PLATES". The default directory for this will be "C:\CELLOMICS\DATA\PLATES". After the creation of the database by the ArrayScan Software, it will then be *SPOOLED* to the server to a subdirectory of the data directory named "PLATES". The default directory for this will be "F:\CELLOMICS\DATA\PLATES". After a few days the database files will be archived and available only by a de-archiving process which can return the database files to the network directory. The Plates table on the server will contain the location of the database to allow the system to locate the data.

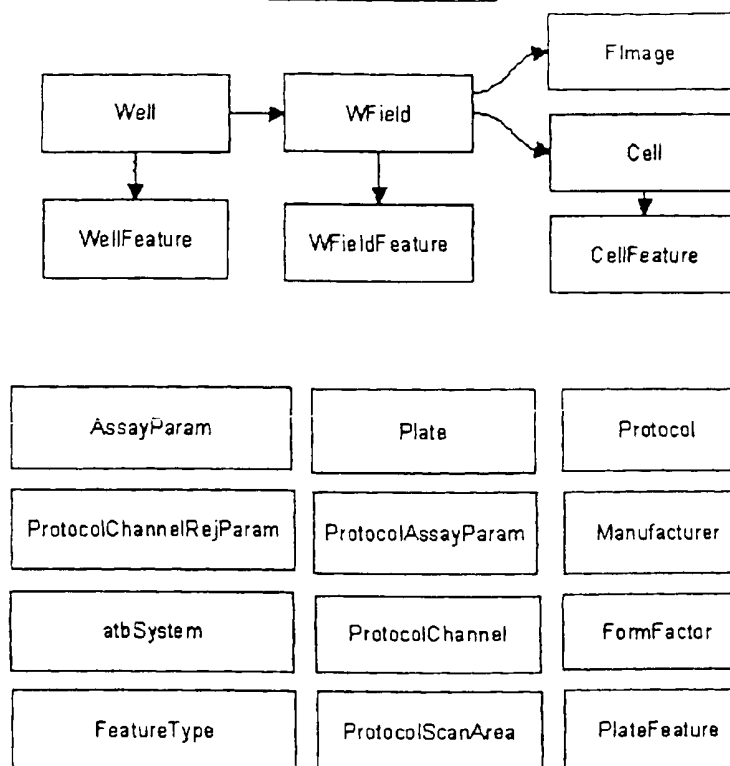
Image files will migrate the same as plate database files, except that they will move to and from subdirectories named "DATA\IMAGES".

5.11. System and Summary Database Tables



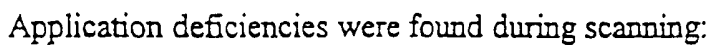
5.12. *Detail Database Tables*

<<Plate.MDB>> Tables



The Specific Plate.MDB for each plate run will consist of the tables to hold the Plate Detail Data and a copy of the tables used for the Plate Results and Configuration. The reason for having a copy of the System.MDB tables in the Plate.MDB is that the Plate.MDB can then be archived and copied to another system for review.

SECRET



☐ Page(s) _____ of _____ were not present
for scanning. (Document title)

☐ Scanned copy is best available.